

<https://helda.helsinki.fi>

Performance profiling as an intelligence-led approach to antidoping in sports

Hopker, James

2020-03

Hopker , J , Griffin , J , Brookhouse , J , Peters , J , Schumacher , Y O & Iljukov , S 2020 , ' Performance profiling as an intelligence-led approach to antidoping in sports ' , Drug Testing and Analysis , vol. 12 , no. 3 , pp. 402-409 . <https://doi.org/10.1002/dta.2748>

<http://hdl.handle.net/10138/328948>

<https://doi.org/10.1002/dta.2748>

unspecified

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Title: Performance profiling as an intelligence-led approach to anti-doping in sports

Authors: James Hopker^a, Jim Griffin^b, James Brookhouse^a, John Peters^c, Yorck Olaf Schumacher^d Sergei Iljukov^{e,f}

^aSchool of Sport and Exercise Sciences, University of Kent, UK.

^bDepartment of Statistical Science, University College London, UK

^cBlue Cat Technical Ltd, West Sussex, UK

^dAspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar

^eKIHU – Research Institute for Olympic Sports, Jyväskylä, Finland

^fDepartment of Sports and Exercise Medicine, Clinicum, University of Helsinki, Finland

Running Head: Performance Profiling and Anti-doping

Corresponding Author:

Dr. James Hopker

School of Sport and Exercise Sciences

University of Kent

Kent

UK

Tel: +44 1634 888814

Email: j.g.hopker@kent.ac.uk

1. Introduction

The current prevalence of doping in elite sport is unknown. Existing data from the World Anti-Doping Agency (WADA) suggests that the number of adverse analytical findings (AAF) was 1.43% in 2017 (4,596 AAFs from 322,050 samples¹). Moreover, data from the recently formed Athletic Integrity Unit (AIU) place the number of anti-doping rule violations (ADRV) at 65 ADRVs from 8,489 samples (<1%)². However, these figures appear at odds to research studies using anonymous athlete self-reports that put the estimate much higher at between 14.0% and 57.1%,³ with between 3.1% and 26.0% of dopers reporting a lifetime prevalence.⁴ Therefore, there is clear discrepancy between the number of anti-doping rule violations (ADRV) and the estimated prevalence of doping. This is despite increases in both the financial resources being allocated to the fight against doping in sports, and number of samples analysed (7.1% increase in the overall number of samples analysed: 300,565 in 2016 to 322,050 in 2017⁵).

Consequently, questions can be raised about the efficiency of current anti-doping policy and testing strategies, and whether other types of data are required. In turn, this targeted approach to anti-doping would enhance the ability of anti-doping authorities to make more informed decisions on assigning athletes to registered testing pools, and target-testing individuals, and ultimately allow a more efficient distribution of anti-doping testing resources.

One key piece of information available to anti-doping authorities is the performance of the athlete, which is currently seldom used in deciding testing strategy. As the primary reason for doping is improvement of athletic performance, it is reasonable to suggest

that monitoring an individual's competition results on a longitudinal basis may reveal suspicious performance improvements. For example, Schumacher & Pottgeisser⁶ have previously demonstrated that yearly world best performances increase with the emergence of new potent doping agents, such as anabolic steroids or EPO, and decrease as their effects become detectable. These findings raise the possibility of monitoring athlete performance in order to inform decisions on anti-doping testing. The main objective of such an "athlete performance passport" (APP) would therefore be to distinguish between expected changes in sporting performances (for example, due to career progression) and disproportionate improvements. The use of performance data in anti-doping is likely to be complementary to analysis of physiological biomarkers for several reasons. The consequences of doping agents manifest themselves at the time of competition, whereas the biomarker itself may only be detectable for a short period of time well in advance of the actual event.⁷ Performance parameters may be affected by subversive practices that are used in an attempt to avoid traditional anti-doping methods. Finally, athletic performance appears relatively consistent and predictable over time⁸ compared to the potential variability in biomarkers used within current anti-doping tests.⁹ In this regard, the identification of unexpected improvements in an athlete's performance could be used as a trigger for their closer scrutiny via a targeted anti-doping testing program.¹⁰ The feasibility of using performance markers for anti-doping is clearest in so-called centimeter, grams, second sports in which competition settings are relatively standardised and the outcome is a clearly measured parameter. Therefore, the athlete's performance potential is predominantly determined by their specific physiological characteristics which are potentially modifiable by doping agents.

However, a key concern with the use of athlete sporting performance data in this context is distinguishing between a physiological increase in performance caused by factors including training & tapering, maturation, seasonal and random variation, from an "unphysiological" improvement caused by doping.¹¹ Analysis of "big data" contained within performance repositories has the potential to increase our understanding of athlete performance and how it evolves over an athlete's career¹² in order to establish *normal* or *expected* rates of progression. Thus, a retrospective trawl of existing data repositories might be used to establish typical performance trends within a large cohort of athletes, and in turn identify career progression that falls outside these normal or expected rates, which may be indicative of doping.^{7,13}

The aim of this case report is to demonstrate the potential for mathematical modelling of individual career trajectories (i.e. the relationship between age and performance) to identify characteristics of performance evolution, which are able to distinguish athletes who have previously been convicted of doping, from others who are presumed clean. In this work, we acknowledge that monitoring athlete performance and identifying changes obviously does not *prove* doping, however extreme changes that are in excess of what is predicted based upon the population changes, may be sufficient to raise the level of suspicion of an athlete. In our analysis, we used a Bayesian spline model¹⁴ to flexibly model the trajectory of each athlete's average performance over their career across a range of athletics disciplines. In the next section, we present our model along with details of how we account for the potential confounding effects of absolute performance level, stage of the athlete within their career, and ageing. Applications of the model

across both track and field disciplines of athletes are presented, and finally we discuss the potential ramifications for current anti-doping policy and practice.

Methodology

Following Institutional ethical approval (Prop_72_2017_18) athletes' data to be included within the model was obtained from the IAAF results database (www.iaaf.org/results).

Specifically, the results of male and female shot put competitions from 2001 to 2018, inclusive, were collected; totaling 1,115 athletes and 56,000 competition results.

Sampled data included the athlete name, IAAF ID number, date of birth, sex, country of birth, event, performance result, date of performance, and finishing position. Doping violations for athletes were obtained from a global database (www.tilastopaja.eu), which is used by a number of sport governing bodies/organizations to track the historic athlete doping violations and performance progression.

Modelling variability in athlete performance

The development of a model for career-long athletic performance trajectories is a key challenge to the use of performance data to identify potential instances of doping. Some issues with this modelling were recently highlighted by Bermon, Metelkina & Rendas,¹⁵ who demonstrated the problems associated with modelling athlete career trajectories caused by differences in individual career length, non-uniformity of distribution of performance events across a calendar year, and the asymmetry of variability in athlete performance (extremely bad performances can result in large deviations from the median, while exceptionally good performances may have smaller departure from the

median). In addition, individual career trajectories show a clear underlying pattern of performance with an initial period of improvement followed by a slow decline due to ageing related factors.

The difference in length of careers and non-uniformity of competition performances was addressed by modelling the relationship between age and career performance using a statistical smoothing method. We did not find that the performances were asymmetric for many athletes in our data set and this effect was not modelled. Our approach first removes the effect of ageing related factors by estimating an average career trajectory using a 5th order polynomial curve applied to the performance of all clean athletes. This allows to define *standardised performance* as the difference between each athlete's actual performances and the average career trajectory curve evaluate at the corresponding age. We find that plotting the standardized performances over time for each athlete shows a clearer pattern of the evolution of performance and its volatility. For this reason, we model standardized performances rather than actual performances.

A standardised performance of zero indicates that an athlete performs at the same level as the average result from the age-matched population, with values above zero representing performances better than the population, and values below zero representing performances that are worse than the population at that age. After adjusting for age-related effects, it would be convenient to use the mean and variance of standardised performance to characterize a range of likely values for each athlete. This could be used to find deviations from the allowable range of standardized

performance in a similar way to the Adaptive Model underlying the Athlete Biological Passport (ABP). However, we find that trends remain in each individual's standardised performance trajectory for both athletes with ADRV and presumed clean athletes. This suggests separately modelling the relationship between age and the standardised performance for each individual using a curving-fitting or smoothing technique. We chose to use a Bayesian spline model. This approach has a number of attractive features. Firstly, it allows us to find ages at which the slope of the standardized performance trajectory changes (which will be called change-points). These points indicate times where performance characteristics are changing, which will, in some cases, be indicative of doping. Secondly, a Bayesian approach allows us to estimate the trend without choosing a number of change-points. This contrasts with other approach which either based on testing for change-points (such as Lombard's test) or model selection procedures. The Bayesian approach also allows us to quantify uncertainty in our estimation of the trend and use this is to quantify uncertainty in predictions about future performance. This is similar to the ABP which also uses a Bayesian method. The ABP assumes that the mean and variance of each blood measurement is constant over time whereas our spline model assumes that the mean standardized performance is changing over time with performance improving or declining linearly between change-points. The change-points indicate where the rate of improvement or decline in standardised performance occur that is in excess of their age-matched counterparts at a similar point in their career. The Bayesian model was fitted using Markov chain Monte Carlo methods. The mathematical basis of the Bayesian Spline model is provided in the accompanying supplementary appendix.

Results - Application of the model

In order to demonstrate our Bayesian approach, we concentrated on male athletes with more than 19 performances. There 405 athletes of which 12 had ADRVs. For simplicity, we will refer to these as doped and clean athletes. The idea of modelling standardised performance plays a key role in our analysis. Figure 1 presents the performance results of two representative doped, and one clean athlete competing in men's shot put events. The top panels present the distances each athlete throws during each competition throughout their career. The bottom panel illustrates the standardised performance for each athlete. That is, how their performance deviates from their age-matched peers across the population of all clean shot put athletes. The standardised performance for the two doped athletes demonstrates a changing career trajectory, along with increases in their performance compared to the expected performance at the times when doping violations were detected. Conversely the clean athletes demonstrate a more consistent trajectory, whether that be gradually increasing or decreasing over their career. This suggests that it is therefore possible to analyse the trajectory of these standardised performances over an athlete's career and statistically identify where changes occur, and the direction of the change. For example, large, rapid increases in an athlete's standardised performance would indicate that their performance (in relation to that of their age and career matched peers) may likely be symptomatic of unnatural enhancements.

INSERT FIGURE 1 HERE

Figure 2 shows the standardised performances trajectories for all athletes in our sample with the mean standardised performance trajectory (shown as a red line). This shows that the standardisation has been successful. The mean standardized performance trajectory is close to zero at most ages with a fairly consistent spread of values over the entire age range.

INSERT FIGURE 2 HERE

Below we present the analysis of representative doped and clean athletes using three plots in order to summarize the results, shown in Figure 3. Column A plots the standardised performance of each athlete as dots, an estimate of the standardised performance trajectory as a solid line with a 95% credible interval shown as a shaded region. The 95% credibility interval indicates a range of values with a 95% probability of containing the true value of the standardised performance trajectory and measures our estimation uncertainty of the trajectory. Column B shows the posterior probability of a change-point at each performance result. For example, the probability of a change-point for Athlete 1 is above 0.1 for a period in this athlete's early 30s. This suggests that there is high probability of a change-point for this Athlete 1 in this period (since the total probability around that age is close to 1). Athlete 3 shows no evidence of a change-point since all values are close to zero. Whereas, Athlete 4 shows evidence of a change-point around 20 and further change-points in their mid- to late-20s. Column C

shows an estimate of the change in standardised performance trajectory as a solid line, with a 95% credible interval shown as a shaded region. This is useful for the understanding the strength of statistical evidence that the slope of standardised performance trajectory is positive or negative at a particular age (where the shaded region is completely above 0 or below 0 respectively). The plots are concentrated on the time periods when an athlete has observed performances.

From the subset of 33 athletes, there were 9 athletes with no evidence of a change point (6 clean, 3 doped). For the doped athletes, there is often a change-point preceding an ADRV (10 of 12 doped athletes). Thus, analysis of the change-point alone may not effectively discriminate between the two groups. However, the change-points indicate ages at which the rate of improvement/decline in standardised performance occur, not the direction of change. Intuitively, we might expect doping to be associated with increases in the rate of improvement and so taking account of the rate of change in standardised performance for each athlete as a function of time is also an important consideration. For doped athletes with change-points preceding their ADRV, there is also often a rapid increase in their performance trajectory, or a sustained period of increase in their standardised performance (e.g. Athlete 1) with the 95% credible interval having positive values. A doped athlete can sometimes show no change point (e.g. Athlete 2) but have increases in standardised performance towards the ADRV date. For non-doped athletes, most have declining standardised performance or small changes (e.g. Athlete 3). However, it is possible to observe two athletes demonstrating

rapid variation/ erratic performance (e.g. Athlete 4), or a profile similar to that of doped athletes (Athlete 5), which might be classed as “suspicious” changes in performance.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

To better understand the overall effect of doping, we compared the standardised performance trajectories of the 37 Scandinavian athletes in our sample and the athletes with positive ADRVs. Scandinavian athletes were selected as a comparator sub-group due to the low reported prevalence of ADRVs across sports.¹⁶ The results are plotted in Figure 4 with a red line showing the average career trajectory. There is a clear pattern of athletes with positive ADRVs showing overall better performance than the Scandinavian athletes after the age of 30. In fact, the slope of the standardized performance shows a clearer difference. Figure 5 shows the proportion of athletes in the two groups for whom the slope of the estimated standardised performance is positive (i.e. their standardised performance is improving). Clearly, the proportion of Scandinavian athletes is decreasing with ages, whereas the proportion of ADRV athletes is increasing. This suggests that ADRV athletes have different features of their standardized performance trajectories.

INSERT FIGURE 5 HERE

Discussion

The development and implementation of the athlete biological passport into the anti-doping practice has acted as a useful deterrent.¹⁷ However, as outlined above, the proportion of adverse analytic findings from anti-doping tests is well below the estimated prevalence of doping with sport. Anti-doping authorities such as WADA and the AIU have therefore both suggested that there is a need for a more intelligence-led approach to anti-doping in order to more effectively target test athletes.¹⁸ The AIU go further to suggest that they aim to be able to risk stratify athletes into high and low risk groups based upon factors such as their athlete biological passport profile and their performance data.¹⁸ The ABP has clear relevance for identification of athletes requiring a great deal of scrutiny and target testing via flagging individuals who have suspicious biomarkers to warrant further investigation, but are not extreme enough for them to transgress an ABP metric. However, the limited time course of a particular drug within an athlete's body system, or window of opportunity to catch inflated blood values may be fairly limited for this approach,¹⁹ even with intensified targeted testing.

By contrast, the performance potential of an athlete tends to be much less volatile. The within-athlete variability of performance for elite athletes over a season in so-called Centimeter-Gram-Second (CGS) sports appears to be relatively small, with for example, the coefficient of variation ranging from 1.1 to 1.4% (90% CI: 1.0–1.6%) in athletic disciplines.⁸ This is because absolute performance level and rate of performance gain are “limited” by known and measurable physiology. Therefore, as demonstrated above, major improvements in performance in relation to that of an athlete's age-matched peers should raise the level of suspicion. For example, athlete 1 (see figure 1)

consistently throws between 19.5 and 21.0 meters throughout most their career (aged between 23 and 34 years), however, towards the end of their career they suddenly start throwing between 21.5 and 22.7 meters (aged 35 years). In practice, this scenario demonstrates an athlete who is barely able to qualify for major competitions, all of a sudden features results that afford a podium finish at World level. Moreover, as this occurs well into the athlete's career when a clear performance trajectory has been established, that level of suspicion is raised further, especially as it coincides with an expected age-related reduction in performance. The athlete in question was subsequently subject to an ARDV for THG. Contrast this profile with the athlete's peer shown by athlete 3 who demonstrates an age-related decline in performance, even though they progress their career around the average for the population (standardised performance and career trajectory are closely distributed around 0). Just as a sudden increase in performance might be a flag of doping, so might sudden decreases, or rapid fluctuations in performance, thus generating an erratic performance profile (e.g. athletes 4 & 5). If the athlete is not able to sustain their doping regime over a prolonged period of time, they may consider attempting to enhance their performance for major competitions, or selection events. Therefore, their performance would be enhanced when doping, but they are not able to maintain the same level of performance indefinitely and regress to their true performance once the doping regimen is finished. In order to avoid an erratic performance profile, the athlete may choose to restrict their competition schedule to include only the major events. However, very rare performances at the elite level may also be suspicious.

It is interesting to note that the same patterns of performance evolution for doped and clean athletes are also evident in other strength- and power-based athletics disciplines, such as sprinting and jumping. However, in long duration athletics events (e.g. middle & long distance running), or sports with high levels of technological, tactical, and environmental influences on performance, there is likely to be an increase the variability of performance results. Such confounding factors would bring a level of unwanted uncertainty into the model prediction, and thus need to be accommodated for by the setting prior distributions in future iterations.

None of the above *proves* doping in previously unsanctioned athletes, although it may be sufficient to raise the level of suspicion of an athlete. What is crucial to this approach is the magnitude, direction, and rate of change in performance in relation to what that athlete has previously been capable of, and that of their age-matched peers. In this way, the anti-doping authorities are able to classify an athlete as high risk or low risk for doping. Combining this information as part of a wider risk prediction model is also likely to further strengthen anti-doping efforts. For example, Marclay, Mangin, Margot & Saugy²⁰ suggest integrating not just analytical chemistry and longitudinal biomarker data, but also information obtained from athlete whereabouts, social media presence, competition level, potential financial rewards, social networks, as well as competitive performance data, into a probability risk type model. The athletes with higher probability of doping risk would therefore be allocated to registered testing pools for closer scrutiny via targeted anti-doping efforts. Moreover, athletes being aware that their performance

is under scrutiny in order to identify unnatural advancements may also serve as a useful deterrent for doping in the first instance.

2. Conclusions

Given the current political and financial climate, there is a need to streamline and refine anti-doping practice with fewer but more targeted testing being conducted. The combination of existing biological data with novel markers, such as athlete performance, has the potential to result in a more intelligence-led and cost-effective anti-doping system. The statistical approach documented in this manuscript demonstrates the potential for analysis of athlete performance data to discriminate between athletes who have been sanctioned for doping, those who are clean, and those who are deemed a suspicious. Therefore, the future development of an athlete performance passport using methods such as those outlined in our manuscript, might be used by anti-doping authorities to discriminate between individuals who are at low or high risk of doping. Further, combining performance data with information provided by the athlete biological passport can provide both a powerful tool in the armory of anti-doping authorities, and a powerful deterrent for athletes considering doping.

Acknowledgements:

The study was funded by a grant from the Partnership for Clean Competition awarded to JH.

References:

1. World Anti-Doping Agency (WADA). Anti-Doping Testing Figures. Available at: https://www.wada-ama.org/sites/default/files/resources/files/2017_anti-doping_testing_figures_en_0.pdf. Accessed March 30, 2019.
2. Athletics Integrity Unity (2019). 2018 Annual Report. Retrieved 8th November 2019 from https://www.athleticsintegrity.org/downloads/pdfs/know-us/en/AIU-Annual-Report-2018-EN_Lite-1.pdf
3. de Hon O, Kuipers H, van Bottenburg M. Prevalence of doping use in elite sports: a review of numbers and methods. *Sports Med* 2015;45: 57–69.
4. Ulrich R, Pope HG, Cleret L, Petroczi A, Nepusz T, Schaffer J, et al. Doping in two elite athletics competitions assessed by randomized-response surveys. *Sports Med* 2018; 48: 211–219.
5. World Anti-Doping Agency (WADA). Annual Report 2017. Available at: <https://www.wada-ama.org/en/resources/finance/annual-report>. Accessed March 30, 2019.
6. Schumacher YO, Pottgiesser T. Performance Profiling: a role for sport science in the fight against doping? *Int J Sports Physiol Perform* 2009;4:129–133.

7. Iljukov S, Berman S, Schumacher Y.O. Application of athletes performance passport for doping control: case report. *Front Physiol* 2018; 9:280.
8. Malcata RM, Hopkins WG. Variability of competitive performance of elite athletes: a systematic review. *Sports Med* 2014;44:1763–1774.
9. Lippi G, Plebani M, Sanchis-Gomar F, Banfi G. Current limitations and future perspectives of the athlete blood passport. *Eur J Appl Physiol* 2010;112: 3693-3694.
10. Hopker J, Schumacher YO, Fedoruk M, Morkeberg J, Berman S, Iljukov S, et al. Athlete performance monitoring in anti-doping. *Front Physiol* 2018;9:232.
11. Iljukov S, Schumacher Y.O. Performance profiling - perspective for anti-doping and beyond. *Front Physiol* 2017;8:1102.
12. Passfield L, Hopker JG. A mine of information: Can sports analytics provide wisdom for your data? *Int J Sports Physiol Perform* 2017;12:851-855.
13. Montagna S, Hopker, J. A Bayesian approach for the use of athlete performance data within anti-doping. *Front Physiol* 2018;9:884.

14. Densson DGT, Holmes CC, Mallick BK, Smith AFM. *Bayesian Methods for Nonlinear Classification and Regression*. Chichester: Wiley; 2002.
15. Bermon S, Metelkina A, Rendas MJ. Comparison of parametric and non-parametric population modelling of sport performances. *Proc Eur Signal Process Conf EUSIPCO* 2018: 306-310.
16. World Anti-Doping Agency (WADA). Anti-doping rule violations (ADRVs) report. Available at: <https://www.wada-ama.org/en/resources/general-anti-doping-information/anti-doping-rule-violations-adrvs-report>. Accessed March 30, 2019.
17. Zorzoli M, Pipe A, Garnier PY, Vouillamoz M, Dvorak J. (2014). Practical experience with the implementation of an athlete's biological profile in athletics, cycling, football and swimming. *Br J Sports Med* 2014;48: 862–866.
18. Athletics Integrity Unit. Athletics Integrity Unit and Abbot World Marathon Majors announce new intelligence and testing programme in groundbreaking anti-doping initiative. Available at: <https://www.athleticsintegrity.org/downloads/pdfs/knowledge-centre/en/AIU-AWMM-PARTNERSHIP-ANNOUNCEMENT-Press-Release.pdf>. Accessed March 30, 2019.

19. Schumacher YO, Saugy M, Pottgiesser T, Robinson N. Detection of EPO doping and blood doping: the haematological module of the athlete biological passport. *Drug Test Anal* 2012;4:846-853.

20. Marclay F, Mangin P, Margot P, Saugy M. Perspectives for forensic intelligence in anti-doping: thinking outside of the box. *Forensic Sci Int* 2013;229:133–144.

Figure Legends:

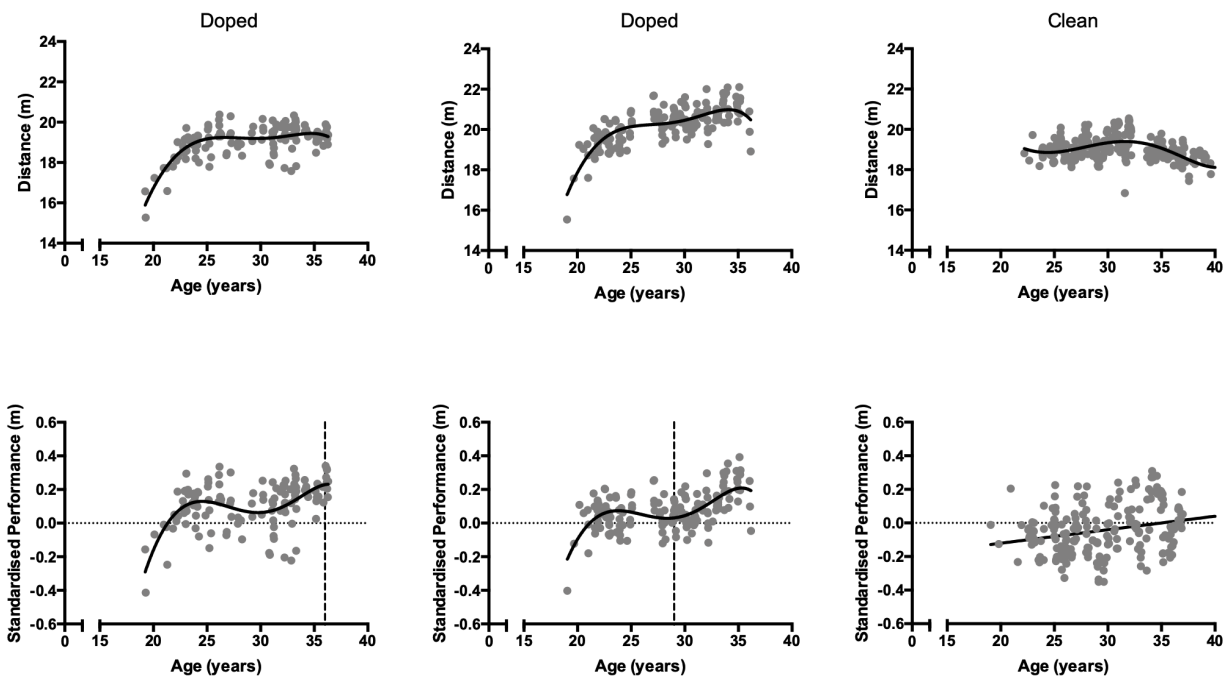


Figure 1: Standardised performance plots of three representative shot put athletes.

Two athletes have previously been sanctioned for doping, and one is presumed clean.

The top panels illustrate the absolute performance of each athletes, the bottom panel presents their standardised performance. Solid lines represent the line of best fit through the standardised performance residuals. Vertical dashed line represents point of sanction. Standardised performance plots distributed around 0 suggest the career evolution of an athlete is close to that expected from the population average. Larger values suggest that performance is greater, with negative values suggesting the performance is lower, than their age matched peers.

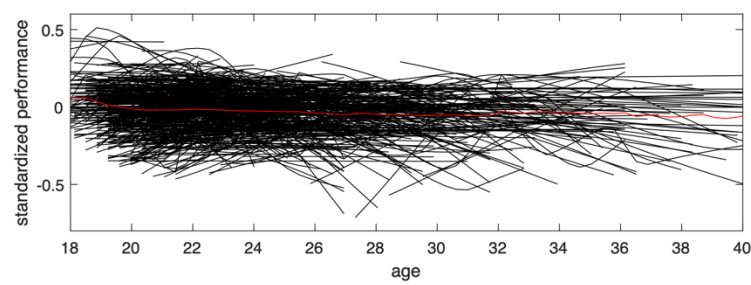


Figure 2: Standardised Performance Trajectories for all athletes in the study. The red line indicates the average Standardised Performance Trajectory.

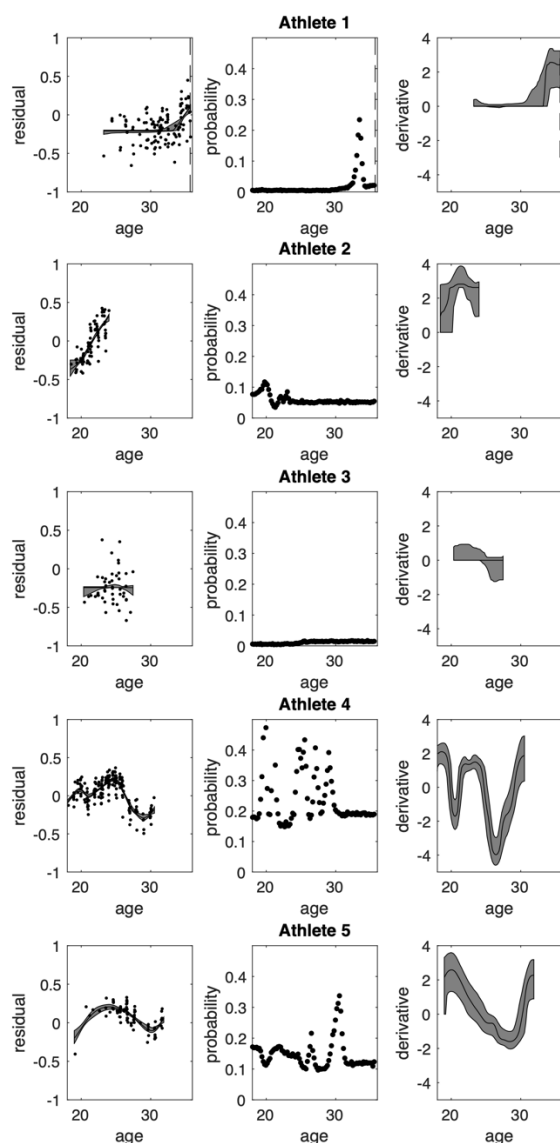


Figure 3: Standardised Performance Trajectories for five representative male shot put athletes from doped and clean populations. Column A = Standardised Performance. Column B = Change point analysis. Column C = Standardised career trajectory. Vertical line indicates timing of ADRV test in doped athletes.

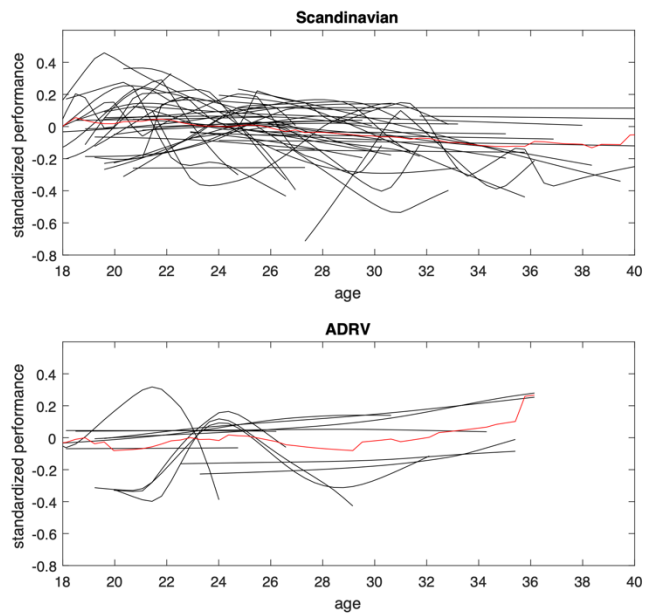


Figure 4: Standardised Performance Trajectories for Scandinavian athletes and athletes with ADRV's. The red line indicates the average Standardised Performance Trajectory for each group.

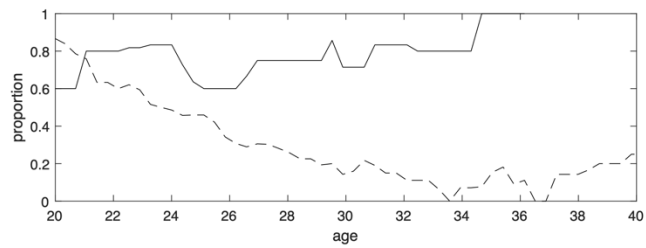


Figure 5: Proportion of Scandinavian athletes (dashed line) and athletes with ADRV's (solid line) with a positive slope of their standardized performance trajectory at different ages.

Technical Appendix to “Performance profiling as an intelligence-led approach to anti-doping in sports”

James Hopker, Jim Griffin, James Brookhouse, John Peters, Yorck Olaf
Schumacher and Sergei Iljukov

This technical appendix provides more details about the model fitted to each athlete’s standardized performances.

The Model

We assume that there is a sample of m athletes and that the i -th athlete has n_i observed standardized performances $y_{i,1}, \dots, y_{i,n_i}$ which occurred at ages $t_{i,1}, \dots, t_{i,n_i}$. The model for the standardized performances of the i -th athlete assumes that:

$$y_{i,j} = f_i(t_{i,j}) + \varepsilon_{i,j}, \quad j = 1, \dots, n_i$$

where $f_i(t)$ is the mean standardized performance of the i -th athlete at age t and $\varepsilon_{i,j} \sim N(0, \sigma_i^2)$ are independent deviations of the observed standardised performance from mean standardized performance. We use a spline model for $f_i(t)$,

$$f_i(t) = \alpha_i + \sum_{k=1}^K C_{i,k} b_{i,k} (t - m_k)^+$$

where $(x)^+ = \max\{0, x\}$ and m_1, \dots, m_K are pre-specified knots (which are chosen to be equally spaced between the smallest and largest age observed with a performance for any athlete), $C_{i,k}$ is 0 or 1 and $b_{i,k}$ are parameters. We refer to $C_{i,k}$ as change points and the slope of the trajectory for the i -th athlete, $f_i(t)$, changes at age m_k if $C_{i,k} = 1$. This

indicates that the rate at which standardised performance is improving or declining has changed for the i -th athlete at that age. The parameter α_i controls the overall level of performance (after standardization). If the i -th athlete's performances are above the average for their peers, α_i will be positive. It is also useful to calculate the derivative of $f_i(t)$ which is denoted

$$\beta_i(t) = \sum_{k=1}^K C_{i,k} b_{i,k} I(t > m_k)$$

where $I(t > m_k) = 1$ if $t > m_k$ and 0 otherwise. The function $\beta_i(t)$ measure the rate of change in standardised performance at time t . If $\beta_i(t)$ is positive, standardized performance is improving and so (unstandardized) performance is improving relative to the average career trajectory, and, if $\beta_i(t)$ is negative, performance is similarly declining. The size of $\beta_i(t)$ measures the rate at which standardised performance is improving or declining.

Bayesian Inference

We refer the interested reader to Denison, Holmes, Mallick & Smith (2002) for further explanation of Bayesian methods and their application to the spline model. The unknown parameters in the model for the i -th athlete are $\alpha_i, C_{i,1}, \dots, C_{i,K}, b_{i,1}, \dots, b_{i,K}$ and σ_i^2 and we choose the following prior distributions. The prior for $C_{i,k}$ is $p(C_{i,k} = 1) = \pi$ and is $p(C_{i,k} = 0) = 1 - \pi$. This implies that prior mean number of change points for the i -th athlete is $K\pi$. We use $\pi=3/K$ which implies that there are three change-points over an athlete's career on average. We use the standard choice $b_{i,k} \sim N(0, \sigma_i^2 g)$ where the value of g is inferred from the data. We also use the standard vague prior distribution

$p(\alpha_i, \sigma_i^2) \propto \sigma_i^{-2}$. Bayesian inference can be implemented in this model using Markov chain Monte Carlo and we refer the reader to Denison, Holmes, Mallick & Smith (2002) for further details.